Introduction

**CMS** uses a tier-structured GRID based computing fabric to provide computing for the CMS collaboration.

The different computing centers use two GRID middleware implementations:

- LHC Computing GRID (LCG) / Enabling Grids for E-sciencE (EGEE)
- OpenScience GRID (OSG)

The user interacts with the GRID via the CMS Remote Batch Builder (CRAB), which is mainly LCG/EGEE centered.

This talk describes the first integration of OSG capabilities into CRAB for the Service Challenge 3 and summaries further plans.
Outline

- CMS computing in the light of LCG/EGEE and OSG
- CRAB introduction
- First OSG implementation
  - Condor-G
  - Additions to CRAB
- Service Challenge 3
- Plans, Summary & Outlook
Data is split into \(~50\) primary datasets (\(~1\) trigger channel)

Datasets are distributed amongst T1 centers (\(8\)-\(16\) per center) and further skimmed to associated T2 centers

Access to distributed datasets using CRAB essential
Middleware within CMS GRID

Europe: LCG / EGEE

USA: OSG
OSG contribution to CMS Tier structure

OSG contribution to CMS tier structure

T1 at FNAL

7 attached T2 sites

<table>
<thead>
<tr>
<th>Site</th>
<th>Processors</th>
<th>Disk (TB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caltech</td>
<td>153</td>
<td>40</td>
</tr>
<tr>
<td>Florida</td>
<td>240+</td>
<td>73</td>
</tr>
<tr>
<td>MIT</td>
<td>(coming soon)</td>
<td>(coming soon)</td>
</tr>
<tr>
<td>Nebraska</td>
<td>256</td>
<td>19</td>
</tr>
<tr>
<td>Purdue</td>
<td>228</td>
<td>~25</td>
</tr>
<tr>
<td>San Diego</td>
<td>228</td>
<td>44.5</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>400</td>
<td>50</td>
</tr>
</tbody>
</table>

For comparison, UAF is ~1000 processors and ~200 TB disk (dCache) -- T2 brings substantial resources to the table.

Today

At least, best guess at the moment; actual disk depends on configuration:

In some cases, four processors/node!

Meant to be 20% of the capacity/complexity required at LHC startup.

Ken Bloom All US CMS Meeting  November 18, 2005
LCG/EGEE and OSG

**LCG / EGEE middleware**

- LHC experiment specific approach using more **higher level tools**
  - access via EGEE tools like **edg-job-submit**
  - utilization of **Resource Broker (RB)**
    - load balancing
    - sandbox for user file input and output to the remote analysis application

**OSG middleware**

- General approach using more **lower level tools** (experiments have to install own tools at OSG sites)
  - access via GLOBUS tools like **globus-job-submit**
  - no RB
  - missing sandbox functionality

➡ **CRAB** build to use RB LCG/EGEE infrastructure and cannot be used directly on OSG sites

➡ **Add** functionality to CRAB to be able to also submit to OSG sites
CMS Remote Batch Builder (CRAB)

CRAB provides the user with a framework to submit user analysis jobs to the CMS tier-structure (T1 and T2 centers)

CRAB takes care of

- User code packing and submission
- Execution on the distant site and status information
- Output retrieval

User interaction with the GRID infrastructure is limited to CRAB usage

More information on CRAB:

- Marco Corvo: CRAB: a tool to enable CMS Distributed Analysis [track: Distributed Data Analysis - 273]
- Daniele Spiga: CRAB usage and jobs-flow Monitoring [track: Distributed Data Analysis - poster]
CRAB user interaction

CRAB splits user interaction into 4 steps

**Job Creation:**
- Resolve requested dataset / MC sample from global CMS service
- Generate list of sites providing access
- Creating jobs in selected job range adding site specific information (jdl, workernode script, POOL catalog retrieval, ...)

**Job Submission**

**Status Check**

**Job Output Retrieval**
OSG middleware is based upon VDT toolkit providing

- **GLOBUS**
- **Condor-G**

First approach: **CONDOR-G** provides:

- **GRID submission** using **GLOBUS** toolkit
- **access** to OSG sites independent of used local batch system
- **sandbox** for insertion and retrieval of files

**Requirements:**

- **OSG site:** none
- **Submitter:** local CONDOR installation with activated CONDOR-G
Implementation: Submitter Side

Main prerequisite:

Locally installed and running Condor scheduler with Condor-G capabilities

First approach: dedicated OSG mode for CRAB

Site list generation excludes non OSG sites

Load balancing of RB not available

first implementation: take first in list

Creation of Condor-G JDL

Use hardcoded information and later GridCat service for site specific information

Changes in workernode script to accommodate OSG implementation

Replacement of LCG / EGEE commands with Condor commands
Submission, Status check and Output retrieval have to be executed from the same scheduler

Submission:

Direct submission to site $\rightarrow$ Order of magnitude faster than RB submission, but: no load balancing, increased load on the Computing Element (CE) of the site

Output retrieval

Condor-G retrieves output automatically $\rightarrow$ no need for dedicated CRAB usage step
Goal of Service Challenge 3 (SC3) in 2005 (11/17/05 - 12/09/05):

- Exercise realistic CMS startup scenario
- Transfer data to tier structure and provide access
- Submit analysis jobs and retrieve output

Tool for analysis part: CRAB

First implementation of OSG additions to CRAB was used during SC3:

- Enabled participation of OSG sites
- 6 OSG T2 participated
**OSG site SC3 participation**

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Caltech</th>
<th>Florida</th>
<th>Nebraska</th>
<th>Purdue</th>
<th>San Diego</th>
<th>Wisconsin</th>
<th>ALL OSG T2 Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>All jobs with zero exit code</td>
<td>8</td>
<td>39</td>
<td>2626</td>
<td>360</td>
<td>733</td>
<td>2496</td>
<td>6262</td>
</tr>
<tr>
<td>All jobs with non-zero exit code</td>
<td>374</td>
<td>61</td>
<td>1172</td>
<td>148</td>
<td>2128</td>
<td>834</td>
<td>4717</td>
</tr>
<tr>
<td>All aborted jobs</td>
<td>409</td>
<td>90</td>
<td>183</td>
<td>629</td>
<td>574</td>
<td>194</td>
<td>2079</td>
</tr>
<tr>
<td>All completed jobs</td>
<td>791</td>
<td>190</td>
<td>3981</td>
<td>1137</td>
<td>3435</td>
<td>3524</td>
<td><strong>13058</strong></td>
</tr>
</tbody>
</table>

13000 analysis jobs have been run against MC samples available at OSG sites transferred during SC3

Each job processed 1000 events and finished well below 8 hours

- zero exit code: job finished with success
- non-zero exit code: specific executable failure exit code
- aborted: problems in GRID submission or output retrieval
Condor-G related failures

- 16% aborted jobs
- 71% of aborted jobs are related to output retrieval errors
- Resulting mostly from application problems
- Remaining errors related to communication problems between submitter and site

### Exit Code Analysis

- **Exit Code Zero**
  - 70.71%
  - The job manager could not stage out a file

- **Exit Code Non-Zero**
  - 29.29%
  - Various errors including:
    - The job manager timed out while waiting for a commit signal
    - The job state file doesn’t exist
    - The standard output/error size is different
    - The user proxy expired (job is still running)
    - The job manager could not stage in a file
    - The scratch directory could not be created
    - The job contact string does not match any which the job manager is handling
    - The job manager could not lock the state lock file

### Condor-G Related Failures

- 16% aborted jobs
  - 16% of aborted jobs are related to output retrieval errors
  - Resulting mostly from application problems

### Table of Globus Errors

<table>
<thead>
<tr>
<th>Globus Error</th>
<th>Description</th>
<th>Percentage [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>An authorization operation failed</td>
<td>12.67</td>
</tr>
<tr>
<td>10</td>
<td>Data transfer to the server failed</td>
<td>0.96</td>
</tr>
<tr>
<td>17</td>
<td>The job failed when the job manager attempted to run it</td>
<td>0.43</td>
</tr>
<tr>
<td>22</td>
<td>The job manager failed to create an internal script argument file</td>
<td>4.38</td>
</tr>
<tr>
<td>24</td>
<td>The job manager detected an invalid script response</td>
<td>0.05</td>
</tr>
<tr>
<td>30</td>
<td>The job manager failed to open the user proxy</td>
<td>1.78</td>
</tr>
<tr>
<td>43</td>
<td>The job manager failed to stage the executable</td>
<td>2.79</td>
</tr>
<tr>
<td>48</td>
<td>The provided RSL could not be properly parsed</td>
<td>0.10</td>
</tr>
<tr>
<td>73</td>
<td>The job manager failed to open stdout</td>
<td>0.19</td>
</tr>
<tr>
<td>74</td>
<td>The job manager failed to open stderr</td>
<td>0.19</td>
</tr>
<tr>
<td>79</td>
<td>Connecting to the job manager failed.</td>
<td>0.24</td>
</tr>
<tr>
<td>93</td>
<td>The gatekeeper failed to find the requested service</td>
<td>0.10</td>
</tr>
<tr>
<td>111</td>
<td>The job manager timed out while waiting for a commit signal</td>
<td>0.14</td>
</tr>
<tr>
<td>121</td>
<td>The job state file doesn’t exist</td>
<td>0.05</td>
</tr>
<tr>
<td>129</td>
<td>The standard output/error size is different</td>
<td>0.10</td>
</tr>
<tr>
<td>131</td>
<td>The user proxy expired (job is still running)</td>
<td>1.01</td>
</tr>
<tr>
<td>135</td>
<td>The job manager could not stage in a file</td>
<td>3.76</td>
</tr>
<tr>
<td>136</td>
<td>The scratch directory could not be created</td>
<td>0.24</td>
</tr>
<tr>
<td>155</td>
<td>The job manager could not stage out a file</td>
<td>70.71</td>
</tr>
<tr>
<td>156</td>
<td>The job contact string does not match any which the job manager is handling</td>
<td>0.05</td>
</tr>
<tr>
<td>158</td>
<td>The job manager could not lock the state lock file</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Plans

OSG additions to CRAB will be finalized in latest CRAB Version 1:

Two submission possibilities will be implemented:

- Condor-G: the presented implementation will be integrated into the new CRAB Version 1

- Resource Broker: OSG sites will report to the LCG / EGEE Resource Broker to enable submission to OSG sites via the RB

Workernode scripts will be adapted to be able to run both on LCG / EGEE sites and OSG sites with “late discovery”
CRAB provides user access to datasets and MC samples distributed worldwide within the CMS tier structure.

The two middlewares, LCG / EGEE and OSG, have both been integrated the first time into the common submission tool.

The OSG enabled CRAB version has been used successfully during the Service Challenge 3.

The GRID / Condor-G related failure rate is very low.

Plans foresee to finalize the OSG implementation into CRAB to enable job submission from a unique interface to OSG sites via the LCG / EGEE resource broker and Condor-G.
**User:**
request to analyze dataset with user code

1. resolve requested dataset into identifier (RefDB)

2. inquire which centers publish requested dataset

3. contact centers and inquire about dataset locally (PubDB)

Jobs are created locally
- on the User’s submission computer
- each job is able to run on all centers from the request list

All the user has to know: which datasamples can I use:

http://cmsdoc.cern.ch/cms/production/www/PubDB/GetPublishedCollectionInfoFromRefDB.mod.php
User’s submitter
- providing created jobs to RB (ship user code to GRID)
- checking status of jobs
- retrieving output (retrieve user code output)

Resource Broker (RB)
- brokers job between requested centers
- provides input and output file handling